

РАЗРАБОТКА СИСТЕМЫ АВТОМАТИЧЕСКОГО РАСПОЗНОВАНИЯ ДОРОЖНЫХ ЗНАКОВ НА ВИДЕО

Кузнецов И.Е.

студент

Научный руководитель И.А. Канаева, ассистент ОИТ

г. Томск, НИ ТПУ, iek7@tpu.ru

Аннотация

В статье представлен алгоритм автоматического распознавания дорожных знаков на видеопоследовательностях. Алгоритм включает в себя детектирование знаков, трекинг, классификацию и извлечение дополнительной информации. Особое внимание уделено подготовке данных, борьбе с дисбалансом классов, алгоритму трекинга, а также фильтрации ложных срабатываний детектора. Достигнута высокая точность классификации (98 %).

Ключевые слова: дорожные знаки, компьютерное зрение, детектирование, трекинг, классификация, RTSD, YOLOv8, SORT, ResNet

Введение

Современная транспортная система предъявляет высокие требования к безопасности и эффективности дорожного движения. Важнейшим элементом обеспечения этих требований является своевременный и точный учет состояния дорожной инфраструктуры, в частности, дорожных знаков. Они несут критически важную информацию для водителей. Традиционные методы паспортизации дорожных знаков, основанные на просмотре видеоматериалов дорожных съемок и ручном переносе данных в базу данных, являются крайне трудоемкими и требуют значительных временных и финансовых затрат.

Задачи автоматического распознавания дорожных знаков решаются уже достаточно давно, и существуют системы, способные выполнять эту функцию, в том числе в режиме реального времени. Однако большинство таких реализаций ориентированы на системы помощи водителю (ADAS), где приоритетом является мгновенное распознавание ограниченного набора критически важных знаков (например, ограничения скорости, стоп, уступи дорогу). Для задач же паспортизации дорожной инфраструктуры предъявляются иные требования: необходимо с высокой точностью распознавать максимально широкий спектр знаков, включая все стандартные типы, их вариации и знаки индивидуального проектирования. Согласно ГОСТ Р 52290-2004 [1], общее число различных дорожных знаков и их разновидностей превышает 300, что ставит перед системами паспортизации сложную задачу по охвату классов.

В этом контексте, целью данной работы является разработка комплексной системы автоматического распознавания дорожных знаков на видеопоследовательностях, адаптированной именно для задач паспортизации, обеспечивающей высокую точность детектирования и классификации широкого спектра знаков и извлекающей расширенный набор характеристик.

Разработанная система обладает рядом отличительных особенностей, определяющих её новизну и прикладную значимость:

1. *Расширенный охват классов:* способность распознавать 223 класса знаков, включая знаки индивидуального проектирования, что критично для полноты паспортизации и превышает возможности типичных ADAS и аналогичных систем.

2. *Оптимизированная скорость обработки для паспортизации:* обработка 10 км дороги занимает 10-25 минут на типовой видеокarte в зависимости от плотности знаков. Известные аналоги могут требовать около 40 минут на тот же объем. Скорость достигается, в том числе, за счет пропуска анализа кадров, на которых детектор не обнаружил знаков.

3. *Продвинутый трекинг*: применение оригинальной гибридной метрики (Centroid+GIoU) в алгоритме SORT [4], специально разработанной для стабильного отслеживания дорожных знаков.

4. *Комплексность извлечения атрибутов*: автоматическое определение не только типа и текста знака, но и его дополнительных параметров, что важно для паспортизации объектов дорожной инфраструктуры.

Основная часть

Система обрабатывает видеоданные из файлов sff (метаданные о дороге: код, название, направление, километраж) и dat (видеокадры, время, пикетаж). Пикетаж представляет собой разметку в виде километровых отметок с начала дороги, позволяющую точно определить местоположение объектов.

В системе используется следующий алгоритм обработки видео:

1. *Чтение данных*: последовательное считывание видеокадров и пикетов из dat файла.
2. *Детектирование*: обнаружение потенциальных дорожных знаков и определение их ограничивающих рамок на каждом кадре с помощью нейронной сети.
3. *Трекинг*: связывание обнаружений одного и того же знака на последовательных кадрах и выбор кадра последнего четкого появления знака для следующих шагов анализа.
4. *Классификация*: определение кода знака с использованием нейронной сети.
5. *Извлечение информации*: распознавание текста на знаке, определение расположения, расстояния, наличия желтой пленки, временного статуса и порядка знаков на одной опоре.
6. *Запись результатов*: сохранение всей собранной информации о распознанных знаках в структурированный CSV-файл.

Подготовка данных. Качество и объем обучающих данных являются критически важными факторами для построения эффективных моделей глубокого обучения. В качестве отправной точки был использован открытый набор данных RTSD (Russian Traffic Sign Dataset) [2], включающий около 50 000 изображений для обучения и 5 000 для тестирования. Однако разметка RTSD охватывала только 155 классов стандартных дорожных знаков, что недостаточно для решения практических задач паспортизации, где часто встречаются знаки индивидуального проектирования (указатели направлений, названия поселений и географических объектов, километровые знаки и т. д.).

Для расширения покрытия классов и повышения робастности моделей были использованы видеоматериалы реальных дорожных съемок. Процесс разметки включал два этапа: предварительную автоматическую разметку с помощью разрабатываемой системы и последующую ручную верификацию и доработку. В ходе ручной доработки добавлялись пропущенные знаки, уточнялись границы и классы, размечались знаки индивидуального проектирования. В результате этой работы к исходному датасету было добавлено около 5500 тщательно размеченных изображений. Общее количество уникальных классов знаков было доведено до 223, что позволило значительно повысить способность системы распознавать более широкий спектр дорожных объектов.

Детектирование. Задача детектирования заключается в локализации всех экземпляров дорожных знаков на входном изображении и определении их ограничивающих рамок. В данной работе для детектирования используется модель YOLOv8, представляющая собой современное развитие семейства однопроходных детекторов YOLO (You Only Look Once) [3]. Первоначально была обучена и протестирована модель YOLOv8 small. Она показала высокую скорость работы, но имела проблемы с обнаружением знаков, расположенных вертикально, а также мелких объектов, таких как таблички дополнительной информации и километровые знаки. Для решения этих проблем была выбрана более крупная и мощная модель YOLOv8 medium. Обучение моделей проводилось на расширенном наборе данных с применением

различных методов аугментации данных, включая случайные геометрические и фотометрические искажения. Обученная модель YOLOv8 medium исправила недостатки младшей модели и продемонстрировала высокую точность и полноту обнаружения всех видов дорожных знаков.

Трекинг. Трекинг объектов на видео необходим для установления соответствия между обнаружениями одного и того же знака на разных кадрах. Это позволяет отфильтровать кратковременные ложные срабатывания детектора, избежать многократного учета одного и того же знака и выбрать наиболее репрезентативный кадр (обычно последнее четкое появление) для последующей классификации и анализа. В системе используется алгоритм SORT (Simple Online and Realtime Tracking) [4]. Он работает в два этапа: предсказание положения треков на текущем кадре с помощью фильтра Калмана и ассоциация предсказанных положений с новыми детекциями с помощью венгерского алгоритма на основе матрицы схожести.

Стандартный алгоритм SORT использует метрику IoU (Intersection over Union) для вычисления схожести. Однако IoU показал недостаточную робастность: при значительном смещении объекта между кадрами, например при приближении к мелким километровым знакам, значение IoU могло достигать нуля, приводя к потере трека. Для преодоления этой проблемы был применен центроидный метод [5], вычисляющий евклидово расстояние между центрами предсказанной рамки трека и рамки детекции.

Этот метод более устойчив к изменениям размера рамки, но его использование в чистом виде привело к другой проблеме: после проезда одного знака трекер мог «перескочить» на следующий знак, расположенный позади, если их центры были достаточно близко. Для решения данной проблемы был внедрен гибридный подход, комбинирующий центроидное расстояние и метрику GIoU (Generalized IoU) [6]. GIoU учитывает не только область пересечения, но и форму и взаимное расположение рамок.

Матрица схожести в текущей версии алгоритма для венгерского алгоритма вычисляется как взвешенная сумма нормализованного центроидного расстояния и GIoU. Этот гибридный подход позволил сохранить устойчивость центроидного метода при изменении масштаба и одновременно предотвратить «перескоки» трекера благодаря учету формы и перекрытия рамок с помощью GIoU, что качественно улучшило стабильность отслеживания.

Классификация. После того как знак был успешно отслежен, изображение, соответствующее его последнему появлению, вырезается и подается на вход классификационной нейронной сети для определения его кода по ГОСТ Р 52290-2004 [1]. Исходный набор данных RTSD страдал от сильного дисбаланса классов (например, больше 24000 примеров для знака пешеходный переход против менее 10 для некоторых редких знаков). Для борьбы с этим применялись стратегии over-sampling для редких классов (дублирование с интенсивной аугментацией до минимум 30 экземпляров) и under-sampling для доминирующих классов (ограничение сверху 1000 случайно выбранными примерами).

Было проведено несколько экспериментов с различными архитектурами и подходами к обучению: использование ResNet50 [7] с дообучением только последнего слоя показало относительно низкую точность. Применение ResNet18 с полным fine-tuning (настройкой всех слоев) позволило значительно улучшить результат.

Наблюдения показали, что детектор иногда срабатывал на объектах, не являющихся дорожными знаками (рекламные щиты, обратные стороны знаков, элементы зданий и пейзажа). Для фильтрации таких ложных детекций на этапе классификации в обучающий набор были добавлены два новых класса: «реклама» и «прочее». Использование более мощной модели ResNet34 с полным fine-tuning в сочетании с добавлением этих фильтрующих классов позволило достичь наивысшей точности классификации.

Извлечение дополнительной информации. Помимо кода знака, система извлекает ряд дополнительных атрибутов знака из изображения. Распознавание текста на знаках

выполняется библиотекой EasyOCR [8]. Также система определяет следующие характеристики: расположение знака относительно дороги, расстояние до знака, наличие желтой светоотражающей пленки, порядок знака на стойке, а также определяется временный это знак или постоянный. Расположение знака определяется на основе его координат и параметров камеры. Расстояние до знака вычисляется с использованием размеров знака на изображении, известных размеров знака в реальном мире и параметров камеры. Наличие желтой пленки и временный статус определяется по цветовым характеристикам изображения знака.

Пользовательский интерфейс. Для обеспечения удобного взаимодействия пользователя с разработанной системой, визуализации результатов распознавания и их последующей верификации был создан графический пользовательский интерфейс (GUI) с использованием библиотеки PyQt6 [9]. Приложение позволяет пользователю выбирать папку с видеоданными для распознавания дорожных знаков, а после открывать результат обработки.

Основное окно приложения (рис. 1) предоставляет комплексное представление данных. Оно включает синхронизированное отображение видеопотоков с нескольких камер, где центральное окно используется для визуализации процесса распознавания (отображение ограничительных рамок, кодов знаков и их иконок), а боковые окна предоставляют дополнительный контекст. Пользователю доступны инструменты для навигации по видео (перемотка, переход по кадрам или пикетажам) и масштабирования изображения.

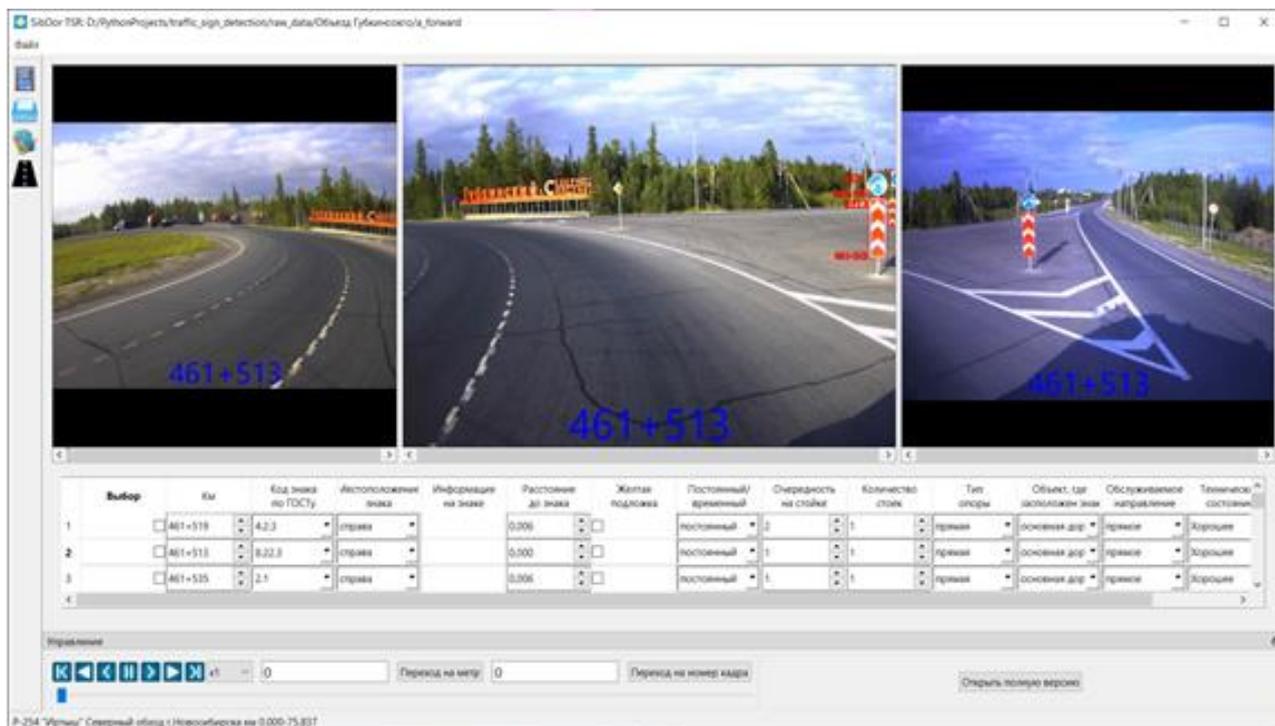


Рис. 1. Основное окно приложения

Ключевым элементом интерфейса является интерактивная таблица, отображающая основную информацию о распознанных дорожных знаках: номер кадра, пикетаж, код знака, расположение, распознанный текст, вычисленное расстояние и другие атрибуты. Таблица позволяет не только просматривать результаты, но и вносить в них корректировки; для полей, связанных со справочниками (например, код знака или его расположение), используются выпадающие списки для удобства ввода. Предусмотрена возможность добавления пропущенных знаков вручную. Результаты обработки и внесенные изменения могут быть сохранены, в том числе с использованием интеграции с базой данных Firebird, используемой для хранения справочников и итоговых данных паспортизации.

Результаты

В данном разделе представлены основные количественные и качественные результаты, достигнутые разработанной системой распознавания дорожных знаков (рис. 2).

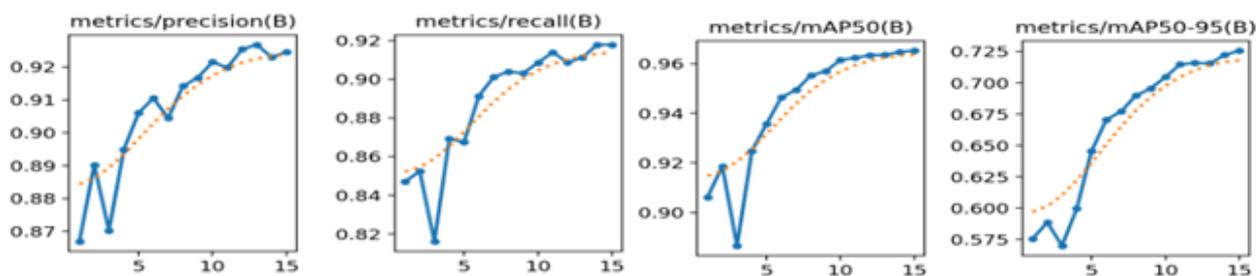


Рис. 2. Графики метрик при обучении детектора YOLOv8m

Производительность финальной модели детектора YOLOv8m, обученной на расширенном наборе данных, иллюстрируется графиками метрик на рисунке 3. Модель достигла показателей точности (precision) 92% и полноты (recall) 91% что свидетельствует о ее способности эффективно локализовывать дорожные знаки на изображениях.

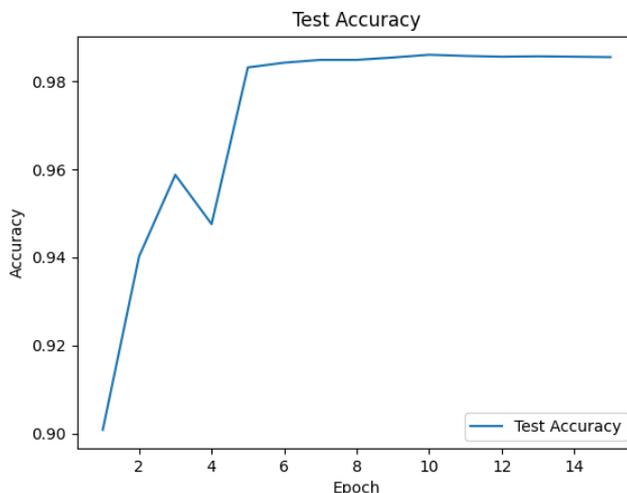


Рис. 3. График изменения точности на тренировочных данных модели ResNet34

Изменение значений точности на тестовых данных в ходе обучения модели классификации ResNet34 представлено на рисунке 3. В таблице 1 представлено значения итоговых метрик трех ранее описанных моделей классификаторов.

Таблица 1. Метрики моделей классификации на тестовой выборке

| Модель | Accuracy | Avg. Precision | Avg. Recall |
|---------------------------|----------|----------------|-------------|
| ResNet50 (последний слой) | 81 % | 79.8 % | 80.1 % |
| ResNet18 | 96 % | 95.6 % | 95.2 % |
| ResNet34 | 98 % | 97.7 % | 97.1 % |

Демонстрация работы алгоритма обработки представлена на рисунке 4, где показаны кадры с одновременным распознаванием множества различных дорожных знаков с корректным определением их кодов. На рисунке 5 показаны примеры распознавания знаков индивидуального проектирования.



Рис. 4. Пример распознавания множества знаков на кадре



Рис. 5. Пример распознавания знаков индивидуального проектирования

Заключение

Представленный алгоритм обработки видео реализует комплексный подход к распознаванию дорожных знаков. Разработанный пользовательский интерфейс обеспечивает удобное взаимодействие с системой. Полученные результаты подтверждают высокую эффективность разработанного подхода и его практическую значимость для автоматизации процессов паспортизации дорожных знаков.

Список использованных источников

1. ГОСТ Р 52290-2004 Технические средства организации дорожного движения. Знаки дорожные. Общие технические требования.
2. Шахуро, В.И. Российская база изображений автодорожных знаков / В.И. Шахуро, А.С. Конушин // Компьютерная оптика. – 2016 – Т. 40, № 2 – С. 294-300.
3. YOLOv8 - Ultralytics YOLO Docs – URL: docs.ultralytics.com/ru/models/yolov8/ (дата обращения: 26.03.2025).
4. Bewley A., Ge Z., Ott L., Ramos F., Upcroft B. Simple Online and Realtime Tracking // 2016 IEEE International Conference on Image Processing (ICIP). Phoenix, AZ: IEEE, 2016. P. 3463-3467.
2. Handle Small/Tiny and Fast/High Speed moving object detection/tracking with stable inference [Электронный ресурс]. – URL: github.com/mikel-brostrom/boxmot/issues/1246 (дата обращения: 27.03.2025).
7. Rezatofighi H., Tsoi N., Gwak J., Sadeghian A., Reid I., Savarese S. Generalized Intersection over Union: A Metric and A Loss for Bounding Box Regression // Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). – Long Beach, CA, USA, 2019. – P. 658-666.
8. He K., Zhang X., Ren S., Sun J. Deep Residual Learning for Image Recognition // 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Las Vegas, NV, USA: IEEE, 2016. P. 770-778.
9. Jaided AI: EasyOCR tutorial [Электронный ресурс]. – URL: github.com/mikel-brostrom/boxmot/issues/1246 (дата обращения: 27.03.2025).
10. Reference Guide-PyQt Documentation v6.8.0 [Электронный ресурс]. – URL: riverbankcomputing.com/static/Docs/PyQt6 (дата обращения: 27.03.2025).